www.ijecs.in
International Journal Of Engineering And Computer Science ISSN: 2319-7242
Volume 6 Issue 1 Jan. 2017, Page No. 20069-20072
Index Copernicus Value (2015): 58.10, DOI: 10.18535/ijecs/v6i1.39

Maximization Influence Flow Using Modified Greedy Algorithm

Jyoti Rani[1], Prof. Ashwani Sethi[2]
[1]Research Scholar M-Tech, Computer Science and Engineering, Guru Kashi University, India
[2]Deputy Director, Guru Kashi University, India
[1]jyotigarg059@gmail.com, [2]@gmail.com

Abstract: The influence maximization problem is of finding a set of seed nodes in social network through which we can maximize the spread of influence. These seed nodes can be used in online marketing so as to maximize the profit. Marketers can get particular group of customers through online market. In this paper we are modifying greedy algorithm so as to maximize results of influence maximization.

Keywords: Influence maximization, social network, spread of influence, independent cascade model, and greedy algorithm.

Introduction

A social network consists of graph of relationships and individual [5]. In this graph relationships are represented by edges and individuals are represented by nodes. If there is an edge between two nodes, it means that there exists a relationship between those two individuals. Main concern of marketing managers is to use this graph and to maximize the spread of information through social networking.

Many online social networking websites like Facebook has become successful because these sites act as effective tools to connect people. These sites are becoming marketing platform which allows flow of information and ideas and have impact on large population in short span of time [6]. Consider an example in which we can understand that why we need to maximize the influence. A small company develops a mobile application let's say an online shopping application. Now, this company wants traffic on this application. But it has limited budget so it can only select small number of initial users and motivates them to use application by providing them special offers (like heavy discounts or by giving them gifts). The company desires that the initial users would start influencing their friends on social networking sites to use the application and their friends would influence their friends' friends and so on, and thus through the word-of-mouth effect a large population in the social network would adopt the application [5, 7]. The problem is who to select as seed users such that influence is on large population. This problem is referred as influence maximization. This problem is of great interest of companies because they can promote their product through this problem.

In this paper we have worked on independent cascade model. In the independent cascade model, the selection procedure from one node to another is based on a probabilistic way; individuals can successfully choose their neighbors with certain probabilities. In this paper we are studying independent cascade model for influence maximization. We are using greedy algorithm. Then we will modify the algorithm so that influence can be maximized and more nodes are get affected through the seed nodes.

Related work

In 2009 [2] Yajun Wang, Siyu Yang study the efficient influence maximization from two complementary directions. One is to improve the original greedy algorithm and its improvement to further reduce its running time, and the other is to design discount heuristics which increases influence spread. Then they evaluate their algorithms and their experimental results show that improved greedy algorithm achieves better running time.

In 2010 [1] Wei chen studied linear threshold model for influence maximization. They proposed influence maximization tailored for linear threshold model. They did simulations to show that proposed algorithm is faster than traditional algorithm.

In 2012 [3] Amit Goyal studied influence maximization from a novel data-based perspective. They introduced a new model, which they call *credit distribution* that explains how influence flows in the network and uses this to estimate expected influence spread. They developed an approximation algorithm for solving the influence maximization problem that provides high accuracy as compared to standard approach.

In 2014 [4] Shengfu Zhou et al says that traditional greedy algorithm is not very efficient for large networks. They proposed a more efficient greedy algorithm. They named it LNG algorithm which works for linear thresh hold model. They performed experiments for large network on their algorithm and showed that the time consuming is very less and spread of influence is better of their proposed algorithm rather than classic greedy algorithm.

Greedy Algorithm

Greedy algorithm is used to maximize the influence of social networking. As the name suggests greedy algorithm will pick the node which have maximum children so that

influence can be maximized. The algorithm for greedy algorithm for maximizing the influence is as given below.

1. Read adjacency matrix

2. totalseeds<- 5

3. seeds <-[25]

4. coverednodes<-[25]

5. Repeat for i from 1 to totalseeds-1

6. iseed=seeds(i)

7. Get children of iseed in array ar.

8. coverednodes=[coverednodes,ar]

9. len=length(ar)

10. Repeat for j from 1 to len

11. el=adjmat(ar(j),:)

12. count(j)=length(el)

13. endloop

14. [xmax ymax]=max(count) //where ymax is index of maximum value

15. seeds(i+1)=ar(ymax)

16. endloop

In above algorithm first adjacency matrix is read. Then total seeds to pick, initial seed and covered nodes (nodes that will be covered or on which nodes influence will be made with seed nodes) are initialized. Then we start a loop in which loop from initial seed and add its children to covered array and then we loop on its children and find a node with maximum children. Then we add that node to seed node and repeat the process until we get total seeds.

Greedy algorithm modified

In our paper we have modified greedy algorithm so that we can maximize the influence or we can say so that we can cover maximum nodes with seed nodes. In this algorithm, instead of choosing node with maximum children we have chosen node with maximum unique nodes. The algorithm for modified greedy algorithm for maximizing the influence is as given below.

1. Read adjacency matrix

2. totalseeds<- 5

3. seeds <-[25]

4. coverednodes<-[25]

5. Repeat for i from 1 to totalseeds-1

6. iseed=seeds(i)

7. Get children of iseed in array ar.

8. coverednodes=[coverednodes,ar]

9. len=length(ar)

10. Repeat for j from 1 to len

11. el=adjmat(ar(j),:)

12. el(ismember(el,coverednodes))=[ ]

13. count(j)=length(el)

14. endloop

15. [xmax ymax]=max(count) //where ymax is index of maximum value

16. seeds(i+1)=ar(ymax)

17. endloop

Similarly, in above algorithm first adjacency matrix is read. Then total seeds to pick, initial seed and covered nodes (nodes that will be covered or on which nodes influence will be made with seed nodes) are initialized. Then we start a loop in which we loop from first seed and find adjacency list of first seed then we added children of seed into covered nodes then for each child (covered node) of seed we find children which are not covered before then we find node with maximum un covered children and add that node into seed node

Experimentation

In our experiment we took a graph of 25 nodes. We make an adjacency list of these nodes. Table 1 shows the adjacency list of 25 nodes.

Table 1: Adjacency list of 25 nodes

Node	Adjacency List
1	2 3 4 10 11
2	1 3 5 6 7
3	1 2 5 7 9 10
4	1 5 8 9 10 11 12 13 16
5	2 3 4 9 10 16 21
6	2 8 9 10 11 13
7	2 3 8 10 12 13 16
8	4 6 7 12 16
9	3 4 5 6 12
10	1 3 4 5 6 7 14 15 24
11	1 4 6 15 16
12	4 7 8 9 13 17
13	4 6 7 12 15 17 18 19
14	10 15 18 19 20
15	10 11 13 14 18 19 20 21 22
16	4 5 7 8 11 19

17	12 13 20 23
18	13 14 15 20 25
19	13 14 15 16
20	14 15 17 18 21
21	5 15 20 22
22	15 21 24 25
23	17 24 25
24	10 22 23
25	18 22 23

After taking adjacency list of 25 nodes greedy algorithm and modified greedy algorithm both are test and outputs of these algorithms are shown in Figure 1 and Figure 2.

```
>> GreedyAlgo
Graph is undirected
Seeds for this graph are:
   25    18    15    10

total number of nodes covered by these seeds

ans =

   18

Nodes covered by these seeds are:
    1     5     8     9    10    11    12    13
 14    15    16    18    19    20    21    22    23    25
```

Figure 1: Output of Greedy algorithm for 4 seeds.

Figure 1 shows that if we took 4 nodes as seed nodes then total nodes that will be covered by these nodes will be 18 in case of greedy algorithm.

```
>> GreedyAlgoModified
Graph is undirected
Seeds for this graph are:
   25    18    13     4

total number of nodes covered by these seeds

ans =

   21

Nodes covered by these seeds are:
  Columns 1 through 19

    1     4     5     6     7     8     9    10    11
 12    13    14    15    16    17    18    19    20    22

  Columns 20 through 21

   23    25
>>
>>
```

Figure 2: Output of modified greedy algorithm for 4 seeds.

Figure 2 shows that if we took 4 nodes as seed nodes then total nodes that will be covered by these nodes will be 21 in case of modified greedy algorithm.

Table 2 shows the tabular form of output for greedy algorithm and modified greedy algorithm in case we select 4 nodes as seed nodes.

Table 2: Nodes covered under greedy and modified greedy algorithm for 4 seed nodes.

Nodes covered under Greedy Algorithm	Nodes covered under modified Greedy Algorithm
18	21

Graphical form of Table 1 is shown in Figure 3.

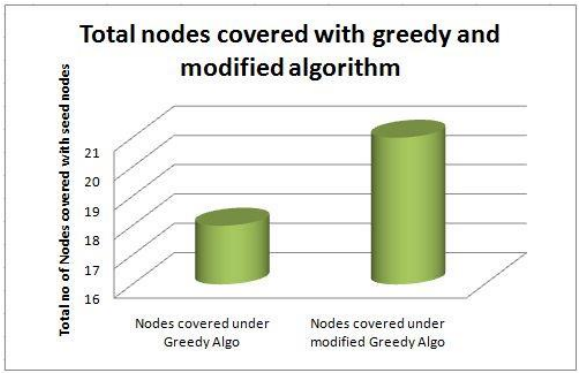

Figure 3: Total nodes covered with and modified greedy algorithm for 4 seed nodes.

In our experimentation we also run list of 25 adjacency nodes after changing total seeds from 4 to 5. Table 3 shows the tabular form of output for greedy algorithm and modified greedy algorithm in case we select 5 nodes as seed nodes.

Table 3: Nodes covered under greedy and modified greedy algorithm for 5 seed nodes.

Nodes covered under Greedy Algorithm	Nodes covered with modified Greedy Algorithm
22	24

Graphical form of Table 2 is shown in Figure 4.

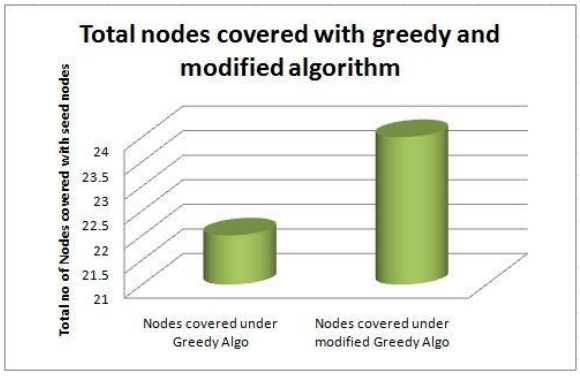

Figure 4: Total nodes covered with and modified greedy algorithm for 5 seed nodes.

Conclusion and Future Scope

The influence maximization is a problem in which we find set of seed nodes such that these seeds can spread maximum influence. In this paper we have worked on greedy algorithm for maximizing the influence. We have first studied greedy algorithm which picks node with maximum number of child. We have modified this algorithm in which we have picked nodes with maximum unique children rather than only maximum children. We have also tested our algorithm which shows that modified algorithm gives better results.

In future further modification will be done on this algorithm to improve results. While selection of next seed we will see children of all seed nodes rather children of selected seed only.

References

[1] Wei Chen, Yifei Yuan and Li Zhang, "Scalable Influence Maximization in Social Networks under the Linear Threshold Model", ICDM '10 Proceedings of the 2010 IEEE International Conference on Data Mining, Pages 88-97, 2010, Washington, DC, USA

[2] Wei Chen, Yajun Wang and Siyu Yang,"Efficient Influence Maximization in Social Networks",ACM SIGKDD international conference on Knowledge discovery and data mining, Pages 199-208, 2009, New York, NY, USA

[3] Amit Goyal, Francesco Bonchi and Laks V. S. Lakshmanan, "A DataBased Approach to Social Influence Maximization", Proceedings of the VLDB Endowment, Vol. 5, No. 1, 2012, pp 73-84

[4] Shengfu Zhou, Kun Yue et.al. ,"An Efficient Algorithm for Influence Maximization under Linear
Threshold Model", Control and Decision Conference (2014 CCDC), The 26th Chinese, 2014, pp 5352 - 5357, Changsha.

[5] Zaixin Lu, Wei Zhang et.al. , "Approximation and Inapproximation for The Influence Maximization Problem in Social Networks under Deterministic Linear Threshold Model", 31st International Conference on Distributed Computing Systems Workshops, 2011, pp. 160-165, Minneapolis, MN.

[6] Goyal, A. and Bonchi, F. and Lakshmanan, L.V.S. and Venkatasubramanian, S., "On minimizing budget and time in influence propagation over social networks," *Social Network Analysis and Mining, 3, 2*, pp. 179–192, 2013.

[7] Jankowski, J. and Michalski, R. and Kazienko, P., "Compensatory Seeding in Networks with Varying Availability of Nodes," in *Proc. of 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining ASONAM 2013*, pp. 1242–1249, 2013.